Afrika Statistika

Afrika Statistika
Vol. 13 (3), 2018, pages 1759 –1777.
DOI: http://dx.doi.org/10.16929/as/1759.132

ISSN 2316-090X

Quadratic error of the conditional hazard function in the local linear estimation for functional data

Torkia Merouan[1], **Boubaker Mechab**[2,*] **and Ibrahim Massim**[3]

Laboratory of Statistics and Stochastic Processes
Department of Probability and Statistics
Djillali Liabes University
Sidi Bel Abbes 22000, Algeria

Received: November 22, 2017. Accepted: September 11, 2018.

Abstract. In this paper we investigate the asymptotic mean square error and the rates of convergence of the estimator based on the local linear method of the conditional hazard function. Under some general conditions, the expressions of the bias and variance are given. The efficiency of our estimator is evaluated through a simulation study. We proved, theoretically and on the scope of a simulation study, that our proposed estimator has better performance than the estimator based on the standard kernel method.

Key words: Nonparametric local linear estimation, conditional hazard function, functional variable, mean squared error.
AMS 2010 Mathematics Subject Classification : 62G05, 62G20.

[*]Corresponding author Boubaker Mechab: mechaboub@yahoo.fr
Torkia Merouan: merouan-to@hotmail.com
Ibrahim Massim : ibrahim.massim6@etu.univ-lorraine.fr

Résumé. Nous étudions dans ce papier, l'estimation non paramétrique de la fonction de hasard conditionnelle basée sur la méthode locale linéaire. Le but est de calculer sous certaines conditions la convergence en moyenne quadratique de notre estimateur, ainsi que les expressions du biais et de la variance de notre estimateur sont données. L'efficacité de notre estimateur est évaluée par une étude de simulation, sur un échantillon fini, qui montre une meilleure performance de l'estimateur introduit par rapport à l'estimateur basée sur la méthode du noyau standard.

1. Introduction

In recent years, the considerable progress in computing power makes it possible to collect and analyze more and more cumbersome data. These large data sets are available primarily through real time monitoring and computers can efficiently deal with such databases.

Many multivariate statistical techniques, concerning parametric models, have been extended to functional data and a good review on this topic can be found in Ramsay and Silverman (2005) or Bosq (2000). Recently, new studies have been carried out in order to propose nonparametric methods taking into account functional data. For a more comprehensive review on this subject the reader is referred to Ferraty and Vieu (2006) and to Ferraty and Vieu (2002) for specialized monographs.

However, it is well known that a local polynomial smoothing procedure has many advantages over the kernel method (see, Fan and Yao (2003) and Fan and Gijbels (1996), etc.). In particular, the former method has better properties, in terms of bias estimation. The local linear smoothing in the functional data setting has been considered by many authors. The first results on the regression function were established in Baìllo and Grané (2009), Boj et al. (2010), Berlinet et al. (2011) and El methni and Rachdi (2011). Other works have been realized on this subject, for example Barrientos-Marin et al. (2010) developed a smoothing local linear estimation of the regression operator for independent data. Moreover, Demongeot et al. (2010) established the almost complete consistency of local linear estimator of the conditional density when the explanatory variable is functional and the observations are i.i.d. The mean squared error of the last estimator was studied by Rachdi et al. (2014). The asymptotic properties (almost complete convergence and convergence in mean square, with rates) of the local linear estimator of the conditional cumulative distribution were established by Demongeot et al. (2014).

This work deals with the functional nonparametric estimation of the hazard and/or the conditional hazard function. Historically, this function was first introduced by Watson and Leadbetter (1964). Since then, several results have been added by many authors. For example Roussas (1989). States that there is extensive literature on nonparametric estimation of the conditional hazard function using a wide variety of methods. This function is important in a variety of fields such as Medicine, Reliability, Survival Analysis or Seismology, etc.

In nonparametric functional framework, the first result has been obtained by Ferraty et al. (2008), who used an approach based on kernel estimations. The authors introduced a kernel estimator of the conditional hazard function and proved some asymptotic properties (with rates) in various situations including censored and/ or dependent variables. Quintela-Del-Río (2008) extended the results of Ferraty et al. (2008). They calculated the bias and variance of these estimates, and established their asymptotic normality. Still while using the Kernel method, Rabhi et al. (2013) determined the asymptotic mean square error of the proposed estimator of the conditional hazard function. In the nonfunctional case, a short overview on nonparametric conditional hazard function estimation can be found in Spierdijk (2008). For functional case, Massim and Mechab (2016) have established the almost complete convergence of the estimator of the conditional hazard function based on the local linear approach.

In the light of what precedes on the importance of the hazard function estimation and the availability of a significant number of advanced and detailed asymptotic results based on the kernel approach, we were interested to find analogous result for the estimator introduced in Massim and Mechab (2016), and next to carry out a thorough comparison with available results.

To achieve this work, we address the described estimator in Massim and Mechab (2016). In this paper, we explicitly determine the mean squared error convergence and compare it to the available result and that obtained through a simulation study.

The remainder of our paper is organized as follows. In section 2, we present our functional model, give basic notations and describe our assumptions. In Section 3. we first state the main theoretical result of the paper about the mean squared convergence in Subsection 3.1 and then, in subsection 3.2, we present the results and we make a comparison with those obtained through simulation study. The proofs are given in Section 4. We conclude the paper by a conclusion and perspective section 5.

2. Description of the Model, Notation and Assumptions

2.1. Model and estimator

Let us consider a sequence $(X_i, Y_i)_{i\geq 1}$ of independent and identically random pair according to the distribution of the pair (X, Y), all of them defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and taking their values in a space $\mathcal{F} \times \mathbb{R}$, where $(\mathcal{F}, d)$ is a semi-metric space.

We suppose that $\mathcal{F} \times \mathbb{R}$ is endowed with the product σ-algebra of the Borel σ-algebras $\mathcal{B}(\mathcal{F})$ and $\mathcal{B}(\mathbb{R})$ on $\mathcal{F}$ and on $\mathbb{R}$ respectively. For a fixed $x \in \mathcal{F}$, we denote by F^x the conditional cumulative distribution function (cdf) of Y given $(X = x)$ and we suppose that F^x is absolutely continuous with respect to the Lebesgue measure with Radon-Nikodym derivative f^x, which is the conditional probability density function (pdf) of Y given $(X = x)$. Accordingly, the conditional hazard function (chf) of Y, given $X = x$, is

$$h^x(y) = \frac{f^x(y)}{1 - F^x(y)}, \ y \in \mathbb{R} \text{ and } F^x(y) < 1. \tag{1}$$

Our main objective is to estimate the conditional hazard function $\widehat{h}^x(\cdot)$ for x fixed, in the form.

$$\widehat{h}^x(y) = \frac{\widehat{f}^x(y)}{1 - \widehat{F}^x(y)}, \ y \in \mathbb{R} \text{ and } \widehat{F}^x(y) < 1. \tag{2}$$

By the fast functional local modeling (cf. Fan (1992)), the conditional cumulative dis- tribution function $F^x(y)$ is estimated as the argmax value of a in the optimization problem, for each $n \geq 1$, the following equation

$$\widehat{F}^x(y) = arg \min_{(a,b)\in\mathbb{R}^2} \sum_{i=1}^{n} \left(H(h_H^{-1}(y - Y_i)) - a - b\beta(X_i, \ x)\right)^2 K(h_K^{-1}\delta(x, X_i)) \tag{3}$$

where $\beta(.,.)$ and $\delta(.,.)$ are locating functions defined from $\mathcal{F} \times \mathcal{F}$ into $\mathbb{R}$, such that:

$$\forall \xi \in \mathcal{F}, \quad \beta(\xi, \xi) = 0 \text{ and } d(\cdot, \cdot) = |\delta(\cdot, \cdot)|.$$

K is a kernel appropriately chosen, H is a distribution function and $h_K = h_{K,n}$(respectively, $h_H = h_{H,n}$) is a sequence of positive real numbers which converges to 0 when $n \to \infty$. Clearly, after direct computations, we get

$$\widehat{F}^x(y) = \frac{\sum\limits_{1\leq i,j\leq n} W_{ij}(x)H({h_H}^{-1}(y - Y_j))}{\sum\limits_{1\leq i,j\leq n} W_{ij}(x)}, \qquad \forall y \in \mathbb{R} \tag{4}$$

with $W_{ij}(x) = \beta_i(\beta_i - \beta_j)K(h_K^{-1}\delta(x, X_i))K(h_K^{-1}\delta(x, X_j))$ and $\beta_i = \beta(X_i, x)$.
Further, the estimator $\widehat{f}^x(y)$ of the density function $f^x(y)$ can be deduced from (4), by

$$\widehat{f}^x(y) = \frac{\sum\limits_{1\leq i,j\leq n} W_{ij}(x)H^{(1)}({h_H}^{-1}(y - Y_j))}{h_H \sum\limits_{1\leq i,j\leq n} W_{ij}(x)}, \qquad \forall y \in \mathbb{R} \tag{5}$$

where $H^{(1)}$ denotes the derivative of H. By putting together Equations 4 and 5, the final form of our estimator (L.M.M.) is: for $n \geq 1, y \in \mathbb{R}$,

$$\widehat{h}^x(y) = \frac{h_H^{-1} \sum\limits_{1\leq i,j\leq n} W_{ij}(x)H'({h_H}^{-1}(y - Y_j))}{\sum\limits_{1\leq i,j\leq n} W_{ij}(x) - \sum\limits_{1\leq i,j\leq n} W_{ij}(x)H({h_H}^{-1}(y - Y_j))}. \tag{6}$$

(See Massim and Mechab (2016)). To be complete, let us remind the conditional hazard function based on the kernel method (K.M.), given for $n \geq 1, y \in \mathbb{R}$ by

$$\widehat{h}^x(y) = \frac{h_H^{-1} \sum_{i=1}^{n} K_i(x) H'({h_H}^{-1}(y - Y_j))}{\sum_{i=1}^{n} K_i(x) - \sum_{i=1}^{n} K_i(x) H({h_H}^{-1}(y - Y_j))}. \tag{7}$$

(See Quintela-Del-Río (2008)). Before we treat the asymptotic theory of the estimator (6) and compare it with that of the estimator (7), we need more notations and clear assumptions given below.

2.2. Notations and assumptions

Let us introduce a set of hypotheses which will be needed to state our main result. Here and below, x (resp. y) will denote a fixed point in $\mathcal{F}$ (resp. $\mathbb{R}$), $\mathcal{N}_x$ (resp. $\mathcal{N}_y$) a fixed neighborhood of a fixed point x (resp. of y) and $\phi_x(r_1, r_2) = \mathbb{P}(r_2 < \delta(X, x) < r_1)$.

(H1) For any $r > 0$, $\phi_x(r) := \phi_x(-r, r) > 0$. There exists a function $\chi_x(\cdot)$ such that

$$\forall t \in (-1, 1), \lim_{h_K \to 0} \frac{\phi_x(th_K, h_K)}{\phi_x(h_K)} = \chi_x(t).$$

(H2)We denote, for any $l \in \{0, 2\}$ and $j = 0, 1$, the functions

$$\psi_{l,j}(x, y) = \frac{\partial^l F^{x^{(j)}}(y)}{\partial y^l} \text{ and } \Psi_{l,j}(s) = \mathbb{E}[\psi_{l,j}(X, y) - \psi_{l,j}(x, y) | \beta(x, X) = s] \tag{8}$$

where $\Psi_{l,j}^{(1)}(0)$ and $\Psi_{l,j}^{(2)}(0)$ of the function $\Psi_{l,j}(\cdot)$ exist and $g^{(k)}$ denotes the k^{th} order derivative of g.

(H3) The functions $\delta(\cdot, \cdot)$ and $\beta(\cdot, \cdot)$ are such that
$\forall z \in \mathcal{F},\ C_1|\delta(x, z)| \leq |\beta(x, z)| \leq C_2|\delta(x, z)|$, with $C_1 > 0, C_2 > 0$,

$$\sup_{u \in B(x,r)} |\beta(u, x) - \delta(x, u)| = o(r)$$

and

$$h_K \int_{B(x,h_K)} \beta(u, x) \mathrm{d}P_X(u) = o\left(\int_{B(x,h_K)} \beta^2(u, x) \mathrm{d}P_X(u) \right)$$

where $B(x, r) = \{x' \in \mathcal{F} : |\delta(x', x)| \leq r\}$.

(H4) The kernel K is a positive, differentiable function which is supported within $(-1, 1)$ satisfies

$$K^2(1) - \int_{-1}^{1} (K^2(u))^{(1)} \chi_x(u) \mathrm{d}u > 0.$$

(H5) The kernel H is a differentiable function which has a bounded first derivative such that

$$\int |t|^2 H^{(1)}(t)\mathrm{d}t < \infty, \int \left(H^{(1)}\right)^2(t)\mathrm{d}t < \infty \text{ and } \int H^{(1)}(t)\mathrm{d}t = 1.$$

(H6) $\exists \alpha < \infty,\ f^x(y) \leq \alpha,\ \forall (x, y) \in \mathcal{F} \times \mathbb{R}$ and

$$\exists 0 < \beta < 1,\ F^x(y) \leq 1 - \beta,\ \forall (x, y) \in \mathcal{F} \times \mathbb{R}.$$

(H7) The bandwidths h_K and h_H satisfy

$$\lim_{n\to\infty} h_K = 0, \quad \lim_{n\to\infty} h_H = 0, \quad \text{and } \lim_{n\to\infty} n h_H^{(j)} \phi_x(h_K) = \infty, \text{ for } j = 0, 1.$$

Some Comments on the assumptions: Assumption (H1) is the concentration property of the explanatory variable in small balls. The function $\chi_x(.)$ plays a fundamental role in all asymptotic study, in particular for the variance term. The condition (H2) is used to control the regularity of the functional space of our model and this is needed to evaluate the bias term of the convergence rates. The assumption (H3) is the same assumption as the assumption (H3) in Rachdi et al. (2014), as introduced in Barrientos-Marin et al. (2010). The hypothesis (H4) and (H5) on the kernels K, H and $H^{(1)}$ are standard conditions in the determination of the quadratic error for functional data. The hypotheses (H6) and (H7) are technical conditions and are also similar to those considered in Ferraty et al. (2008).

3. RESULTS

In this section we are going to state our theoretical results. In the first subsection the proof of our main Theorem 1 is demonstrated in terms of Theorems 2-3 and Lemmas 1-7. The full proofs of all these theoretical results are postponed to Section 4. As a result, we will have the time to focus on the simulation study in the second subsection of the current section.

3.1. Main results: Mean Squared Convergence

Theorem 1. *Under assumptions (H1)-(H7), we obtain*

$$\mathbb{E}\left[\widehat{h}^x(y) - h^x(y)\right]^2 = B_n^2(x, y) + \frac{V_{HK}(x, y)}{n h_H \phi_x(h_K)} + o(h_H^4) + o(h_K^4) + o\left(\frac{1}{n h_H \phi_x(h_K)}\right)$$

where

$$B_n(x,y)=\frac{(B_{f,H}-h^x(y)B_{F,H})h_H^2+(B_{f,K}-h^x(y)B_{F,K})h_K^2}{1-F^x(y)}$$

with

$$\begin{aligned}
B_{f,H}(x,y) &= \tfrac{1}{2}\tfrac{\partial^2 f^x(y)}{\partial y^2}\int t^2H^{(1)}(t)\mathrm{d}t\\
B_{f,K}(x,y) &= \tfrac{1}{2}\Psi_{0,1}^{(2)}(0)\left[\frac{\left(K(1)-\int_{-1}^{1}(u^2K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1)-\int_{-1}^{1}K^{(1)}(u)\chi_x(u)\mathrm{d}u\right)}\right]\\
B_{F,H}(x,y) &= \tfrac{1}{2}\tfrac{\partial^2 F^x(y)}{\partial y^2}\int t^2H^{(1)}(t)\mathrm{d}t\\
B_{F,K}(x,y) &= \tfrac{1}{2}\Psi_{0,0}^{(2)}(0)\left[\frac{\left(K(1)-\int_{-1}^{1}(u^2K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1)-\int_{-1}^{1}K^{(1)}(u)\chi_x(u)\mathrm{d}u\right)}\right]
\end{aligned}$$

and

$$V_{HK}^h(x,y)=\frac{h^x(y)}{(1-F^x(y))}\left[\frac{\left(K^2(1)-\int_{-1}^{1}(K^2(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1)-\int_{-1}^{1}(K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)^2}\right].$$

Comparison remark. It is clear that the bias term is of order of the (L.M.M.) estimator and is given by

$$C_Hh_H^2+C_Kh_K^2. \tag{9}$$

We already know form available literature (see for example, Quintela-Del-Río (2008) and Rabhi et al. (2013)) that bias term of the (K.M.) estimator is

$$C_Hh_H^2+C_Kh_K. \tag{10}$$

Further more, both (LMM) and (KM) estimators have equivalent asymptotic variances functions.

From these two remarks, the (LMM) estimator behaves better that the (KM) estimator since $h_K\to 0$ as $n\to+\infty$.
In subsection 3.2, we will confirm this important result by simulations.

Below we will just show how Theorem 1 is proved as a subsequent result of Theorems 2-3 which are fully proved in Section 4.

Proof of Theorem 1. By using the following decomposition

$$\begin{aligned}
\widehat{h}^x(y)-h^x(y) &= \frac{1}{1-\widehat{F}^x(y)}\left[(\widehat{f}^x(y)-f^x(y))+\frac{f^x(y)}{1-F^x(y)}(\widehat{F}^x(y)-F^x(y))\right]\\
&\leqslant \frac{1}{1-\widehat{F}^x(y)}\left[(\widehat{f}^x(y)-f^x(y))+\frac{\tau}{\beta}(\widehat{F}^x(y)-F^x(y))\right]\\
&\leqslant \left[(\widehat{f}^x(y)-f^x(y))+\frac{\alpha}{\beta}(\widehat{F}^x(y)-F^x(y))\right].
\end{aligned}$$

The proof of Theorem 1 can be deduced from Theorem 2, Theorem 3 and the following result

$$\exists \epsilon > 0 \text{ such that} \sum_{n\in\mathbb{N}} \mathbb{P}(1 - \widehat{F}^x(y) < \epsilon) < \infty. \tag{11}$$

Theorem 2. *Under assumptions (H1)-(H7), we obtain*

$$\begin{aligned}\mathbb{E}\left[\widehat{f}^x(y) - f^x(y)\right]^2 &= B^2_{f,H}(x,y)h_H^4 + B^2_{f,K}(x,y)h_K^4 + \frac{V^f_{HK}(x,y)}{nh_H\phi_x(h_K)} \\ &\quad + o(h_H^4) + o(h_K^4) + o\left(\frac{1}{nh_H\phi_x(h_K)}\right)\end{aligned}$$

where $V^f_{HK}(x,y) = f^x(y)\left[\frac{\left(K^2(1) - \int_{-1}^1 (K^2(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1) - \int_{-1}^1 (K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)^2}\right]\int (H^{(1)}(t))^2 \mathrm{d}t.$

We set

$$\widehat{f}^x_N(y) = \frac{1}{n(n-1)h_H\mathbb{E}[W_{12}(x)]}\sum_{1\le i\neq j\le n} W_{ij}(x)H^{(1)}(h_H^{-1}(y - Y_j))$$

and

$$\widehat{f}_D(x) = \frac{1}{n(n-1)\mathbb{E}[W_{12}(x)]}\sum_{1\le i\neq j\le n} W_{ij}(x)$$

then

$$\widehat{f}^x(y) = \frac{\widehat{f}^x_N(y)}{\widehat{f}_D(x)}.$$

The proof of Theorem 2 can be deduced from the following intermediates results.

Lemma 1. *Under the hypotheses of Theorem 2, we get*

$$\mathbb{E}\left[\widehat{f}^x_N(y)\right] - f^x(y) = B_{f,H}(x,y)h_H^2 + B_{f,K}(x,y)h_K^2 + o(h_H^2) + o(h_K^2).$$

Lemma 2. *Under the hypotheses of Theorem 2, we have*

$$Var\left[\widehat{f}^x_N(y)\right] = \frac{V^f_{HK}(x,y)}{nh_H\phi_x(h_K)} + o\left(\frac{1}{nh_H\phi_x(h_k)}\right).$$

Lemma 3. *Under the hypotheses of Theorem 2, we get*

$$Cov(\widehat{f}^x_N(y), \widehat{f}_D(x)) = O\left(\frac{1}{n\phi_x(h_K)}\right).$$

Lemma 4. *Under the hypotheses of Theorem 2, we have*

$$Var\left[\widehat{f}_D(x)\right] = O\left(\frac{1}{n\phi_x(h_K)}\right).$$

Theorem 3. *Under assumptions (H1)-(H7), we obtain*

$$\begin{aligned}\mathbb{E}\left[\widehat{F}^x(y) - F^x(y)\right]^2 &= B^2_{F,H}(x,y)h_H^4 + B^2_{F,K}(x,y)h_K^4 + \frac{V^F_{HK}(x,y)}{n\phi_x(h_K)} \\ &+o(h_H^4) + o(h_K^4) + o\left(\frac{1}{n\phi_x(h_K)}\right)\end{aligned}$$

where $V^F_{HK}(x,y) = F^x(y)(1-F^x(y))\left[\frac{\left(K^2(1)-\int_{-1}^1 (K^2(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1)-\int_{-1}^1 (K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)^2}\right]$.

We note that

$$\widehat{F}^x(y) = \frac{\widehat{F}^x_N(y)}{\widehat{f}_D(x)}$$

where

$$\widehat{F}^x_N(y) = \frac{1}{n(n-1)\mathbb{E}[W_{12}(x)]}\sum_{1\le i\ne j\le n} W_{ij}(x)H(h_H^{-1}(y-Y_j)).$$

The following lemmas will be useful for proof of Theorem 3.

Lemma 5. *Under the hypotheses of Theorem 3, we get*

$$\mathbb{E}\left[\widehat{F}^x_N(y)\right] - F^x(y) = B_{F,H}(x,y)h_H^2 + B_{F,K}(x,y)h_K^2 + o(h_H^2) + o(h_K^2).$$

Lemma 6. *Under the hypotheses of Theorem 3, we have*

$$Var\left[\widehat{F}^x_N(y)\right] = \frac{V^F_{HK}(x,y)}{n\phi_x(h_K)} + o\left(\frac{1}{n\phi_x(h_k)}\right).$$

Lemma 7. *Under the hypotheses of Theorem 3, we get*

$$Cov(\widehat{F}^x_N(y), \widehat{f}^x_D) = O\left(\frac{1}{n\phi_x(h_K)}\right).$$

3.2. Simulation study on the finite samples

We have already justified, as mentioned in the comparison remark given after the statement of Theorem 1, how the (LMM) estimator given in Formula (6) should behave better that the (KM) estimator given in Formula (7). We are going to illustrate this by a simple simulation experience.

Let us fix a functional regression model,

$$Y_i = m(X_i) + \varepsilon$$

where the random variable ε is normally distributed as $\mathcal{N}(0,1)$ and

$$m(x) = 4\exp\left(\frac{1}{1+\int_0^\pi |x(t)|^2 dt}\right).$$

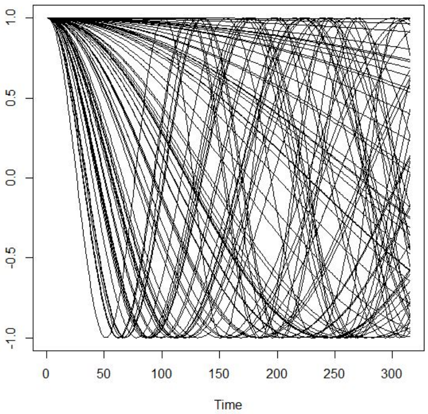

Fig. 1. Curves X_i

The functional variable X is chosen as a real-valued function with support $[0, \pi]$, we generate $n = 100$ functional data (see Figure 1) by:

$$X_i(t) = \cos(W_i(t)), \text{for all } t \in [0, \pi] \text{ et } i = 1, ..., n$$

where the random variables W_i are i.i.d. and follow the normal distribution $\mathcal{N}(0, 1)$. The curves are discretized on the same grid which is composed of 100 equidistant values in $[0, \pi]$.

Based on this data, we generated the (LMM) and the (KM) statistics. First, we compare the two obtained graphs, each of both compared with the true conditional hazard function in Figure 2.

Next, we compare the performance of both estimators by means of the absolute error (AE) defined

$$AE = |\text{true value} - \text{estimated value}|. \tag{12}$$

We report the results of the computations of the AE's in Table 1.

Form the graphs and the tables, we may draw a number of useful comments.

(a) In Figure 1, it can be seen that the (LMM) estimator fits better the chf than the (KM) estimator.

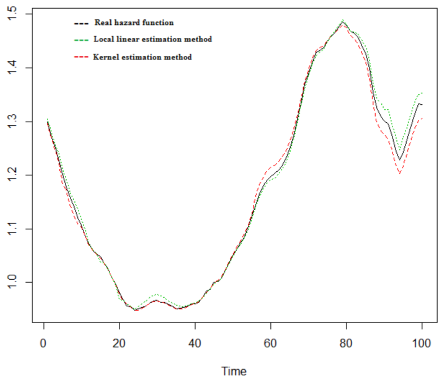

Fig. 2. Comparison between estimation methods

Table 1. Comparison of the AE's

Number of sample	AE (L.L.M)$\times 10^3$	AE (K.M)$\times 10^3$
10	11.9	15.1
20	1.6	1.9
40	0.7	1.3
60	6.2	17.4
80	0.6	4.8
100	19.1	25.4

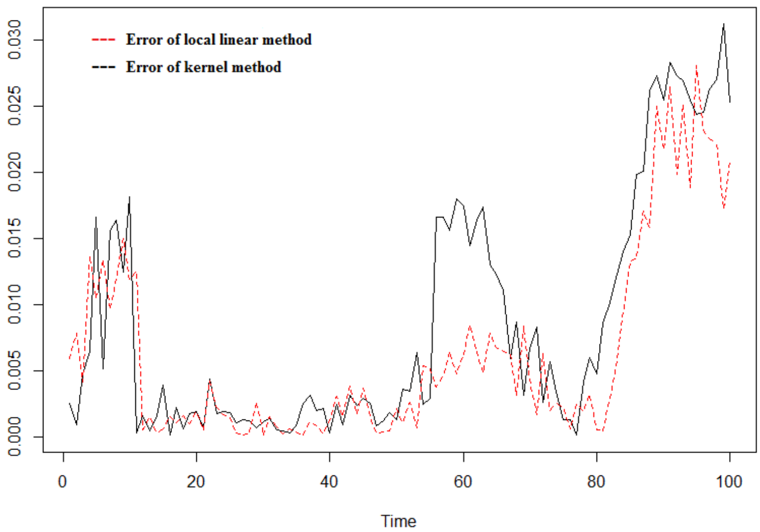

Fig. 3. The AE-errors of both methods

(b) From Table 1, we see that the absolute error for the local linear estimation method, in most cases, is smaller than the absolute error in the kernel estimation method.

(c) The full graphs of the AE'S are illustrated in Figure 3.

As a general conclusion, we may say the (LMM) estimator performance is better than that of the (KM) with respect to the absolute error and the bias for n large so that h_K is small enough to impact the comparison.

4. Proofs

In the proofs below, we will need the following additional notation. C strictly positive generic constant. For all $(i, j) \in \{1, ..., n\}^2$, we have

$$K_i = K(h_K^{-1}\delta(X_i, x)),\ W_{ij} = W_{ij}(x),$$

$$H_j = H(h_H^{-1}(y - Y_j)), H_j^{(1)} = H_j^{(1)}(h_H^{-1}(y - Y_j)).$$

The proofs are organized as follows.
Theorem 2 presents the mean square error of the conditional density estimator. To prove this theorem we need to prove lemmas 1-4. Similarly, to prove Theorem 3 which presents present the mean square error of the conditional distribution estimator, we need to prove lemmas 5-7.

Proof of Theorem 2. We begin by computing the bias and the variance of $\widehat{f}^x(y)$. We have

$$\mathbb{E}\left[\widehat{f}^x(y) - f^x(y)\right]^2 = \left[\mathbb{E}\left[\widehat{f}^x(y)\right] - f^x(y)\right]^2 + Var\left[\widehat{f}^x(y)\right]. \tag{13}$$

By simple calculations, we get

$$\begin{aligned}\widehat{f}^x(y) - f^x(y) &= \left(\widehat{f}_N^x(y) - f^x(y)\right) - \left(\widehat{f}_N^x(y) - \mathbb{E}[\widehat{f}_N^x(y)]\right)\left(\widehat{f}_D(x) - 1\right) \\ &\quad -\mathbb{E}[\widehat{f}_N^x(y)]\left(\widehat{f}_D(x) - 1\right) + \left(\widehat{f}_D(x) - 1\right)^2 \widehat{f}^x(y).\end{aligned}$$

From that fact that $\mathbb{E}[\widehat{f}_D(x)] = 1$, we deduce that:

$$\begin{aligned}\mathbb{E}\left[\widehat{f}^x(y)\right] - f^x(y) &= \left(\mathbb{E}[\widehat{f}_N^x(y)] - f^x(y)\right) - Cov\left(\widehat{f}_N^x(y), \widehat{f}_D(x)\right) \\ &\quad +\mathbb{E}\left[\left(\widehat{f}_D(x) - \mathbb{E}[\widehat{f}_D(x)]\right)^2 \widehat{f}^x(y)\right].\end{aligned}$$

Since the kernel $H^{(1)}$ is bounded, we can bound $\widehat{f}^x(y)$ by a constant $C > 0$, where $\widehat{f}^x(y) \leq C/h_H$. Hence

$$\begin{aligned}\mathbb{E}\left[\widehat{f}^x(y)\right] - f^x(y) &= \left(\mathbb{E}[\widehat{f}_N^x(y)] - f^x(y)\right) - Cov\left(\widehat{f}_N^x(y), \widehat{f}_D(x)\right) \\ &\quad +Var\left[\widehat{f}_D(x)\right] O(h_H^{-1}).\end{aligned}$$

Now, by Bosq and Lecoutre (1987), the variance term in (13) is

$$\begin{aligned} Var\left[\widehat{f}^x(y)\right] &= Var\left[\widehat{f}_N^x(y)\right] - 2\mathbb{E}[\widehat{f}_N^x(y)]Cov\left(\widehat{f}_N^x(y), \widehat{f}_D(x)\right) \\ &+ \left(\mathbb{E}[\widehat{f}_N^x(y)]\right)^2 Var\left(\widehat{f}_D(x)\right) o\left(\frac{1}{nh_H\phi_x(h_K)}\right). \blacksquare \end{aligned}$$

Proof of Lemma 1. We have

$$\mathbb{E}[\widehat{f}_N^x(y)] = \frac{1}{h_H\mathbb{E}[W_{12}]}\mathbb{E}\left[W_{12}\mathbb{E}[H_2^{(1)}|X_2]\right]. \tag{14}$$

By using a Taylor's expansion and under assumption $(H5)$, we get

$$\mathbb{E}[H_2^{(1)}|X_2] = f^{X_2}(y) + \frac{h_H^2}{2}\left(\int t^2H^{(1)}(t)\mathrm{d}t\right)\frac{\partial^2 f^{X_2}(y)}{\partial y^2} + o(h_H^2).$$

The latter can be re-written as

$$\mathbb{E}[H_2^{(1)}|X_2] = \psi_{0,1}(X_2, y) + \frac{h_H^2}{2}\left(\int t^2H^{(1)}(t)dt\right)\psi_{2,1}(X_2, y) + o(h_H^2).$$

Thus, from (14), we obtain

$$\mathbb{E}\left[\widehat{f}_N^x(y)\right] = \frac{1}{E[W_{12}]}\left(\mathbb{E}\left[W_{12}\psi_{0,1}(X_2, y)\right] + \frac{h_H^2}{2}\left(\int t^2H^{(1)}(t)\mathrm{d}t\right)\mathbb{E}\left[W_{12}\psi_{2,1}(X_2, y)\right] + o(h_H^2)\right).$$

Accordingly with to Ferraty et al. (2007), we may shot that for $l \in \{0, 2\}$,

$$\begin{aligned} \mathbb{E}[W_{12}\psi_{l,1}(X_2, y)] &= \psi_{l,1}(x, y)\mathbb{E}[W_{12}] + \mathbb{E}[W_{12}(\psi_{l,1}(X_2, y) - \psi_{l,1}(x, y))] \\ &= \psi_{l,1}(x, y)\mathbb{E}[W_{12}] + \mathbb{E}[W_{12}\mathbb{E}[\psi_{l,1}(X_2, y) - \psi_{l,1}(x, y)|\beta(X_2, x)]] \\ &= \psi_{l,1}(x, y)\mathbb{E}[W_{12}] + \mathbb{E}[W_{12}\Psi_{l,1}(\beta(X_2, x))]. \end{aligned}$$

By Observing that $\Psi_{l,1}(0) = 0$ and $\mathbb{E}\left[\beta(X_2, x)W_{12}\right] = 0$, we get

$$\mathbb{E}\left[W_{12}\psi_{l,1}(X_2, y)\right] = \psi_{l,1}(x, y)\mathbb{E}[W_{12}] + \frac{1}{2}\Psi_{l,1}^{(2)}(0)\mathbb{E}\left[\beta^2(X_2, x)W_{12}\right] + o(\mathbb{E}\left[\beta^2(X_2, x)W_{12}\right]).$$

So,

$$\begin{aligned} \mathbb{E}\left[\widehat{f}_N^x(y)\right] &= f^x(y) + \frac{h_H^2}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int t^2H^{(1)}(t)\mathrm{d}t + o\left(h_H^2\frac{\mathbb{E}\left[\beta^2(X_2, x)W_{12}\right]}{\mathbb{E}[W_{12}]}\right) \\ &+ \Psi_{0,1}^{(1)}(0)\frac{\mathbb{E}\left[\beta^2(X_2, x)W_{12}\right]}{2\mathbb{E}[W_{12}]} + o\left(\frac{\mathbb{E}\left[\beta^2(X_2, x)W_{12}\right]}{\mathbb{E}[W_{12}]}\right). \end{aligned}$$

The two quantities $\mathbb{E}\left[\beta(x, X_2)^2W_{12}\right]$ and $\mathbb{E}[W_{12}]$ are based on the asymptotic evaluation of $\mathbb{E}[K_1^a\beta_1^b]$ (see Rachdi et al. (2014) for more details). To do that, first we treat the case $b = 1$ and $a > 0$. For this case, we use the last part of (H3) and (H4), to get

$$h_K\mathbb{E}[K_1^a\beta_1] = o\left(\int_{B(x,h_K)}\beta^2(u, x)\mathrm{d}P_X(u)\right) = o(h_K^2\phi_x(h_K)).$$

So, we can see that,

$$\mathbb{E}[K_1^a\beta_1] = o(h_K\phi_x(h_K)). \tag{15}$$

On the other hand, for all $b > 1$, and after simplifications of the expressions, we have

$$\mathbb{E}[K_1^a\beta_1^b] = \mathbb{E}[K_1^a\delta^b(x,X)] + o(h_K^b\phi_x(h_K)).$$

Concerning the first term, we write

$$\begin{aligned}
h_K^{-b}\mathbb{E}[K_1^a\delta^b] &= \int v^bK^a(v)\mathrm{d}P_X^{h_K^{-1}\delta(x,X)}(v)\\
&= \int_{-1}^{1}\left[K^a(1) - \int_v^1\left((u^bK^a(u))^{(1)}\right)\mathrm{d}u\right]\mathrm{d}P_X^{h_K^{-1}\delta(x,X)}(v)\\
&= \left(K(1)\phi_x(h_K) - \int_{-1}^{1}(u^bK^a(u))^{(1)}\phi_x(uh_K,h_K)\mathrm{d}u\right)\\
&= \phi_x(h_K)\left(K(1) - \int_{-1}^{1}(u^bK^a(u))^{(1)}\frac{\phi_x(uh_K,h_K)}{\phi_x(h_K)}\mathrm{d}u\right).
\end{aligned}$$

Then, under assumptions $(H1)$, we get

$$\mathbb{E}[K_1^a\beta_1^b] = h_K^b\phi_x(h_K)\left(K(1) - \int_{-1}^{1}(u^bK^a(u))^{(1)}\chi_x(u)\mathrm{d}u\right) + o(h_K^b\phi_x(h_K)). \tag{16}$$

So,

$$\frac{\mathbb{E}[\beta^2(X_2,x)W_{12}]}{\mathbb{E}[W_{12}]} = h_K^2\left(\frac{K(1) - \int_{-1}^{1}(u^2K(u))^{(1)}\chi_x(u)\mathrm{d}u}{K(1) - \int_{-1}^{1}(K^{(1)}(u)\chi_x(u)\mathrm{d}u}\right) + o(h_K^2).$$

Hence,

$$\begin{aligned}
\mathbb{E}\left[\widehat{f}_N^x(y)\right] &= f^x(y) + \frac{h_H^2}{2}\frac{\partial^2 f^x(y)}{\partial y^2}\int t^2H^{(1)}(t)\mathrm{d}t + o(h_H^2)\\
&\quad + \frac{h_K^2}{2}\Psi_{0,1}^{(2)}(0)\frac{\left(K(1) - \int_{-1}^{1}(u^2K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1) - \int_{-1}^{1}K^{(1)}(u)\chi_x(u)\mathrm{d}u\right)} + o(h_K^2). \blacksquare
\end{aligned}$$

Proof of Lemma 2. We have

$$Var\left(\widehat{f}_N^x(y)\right) = \frac{1}{(n(n-1)h_H(\mathbb{E}[W_{12}]))^2}Var\left(\sum_{1\le i\ne j\le n}W_{ij}H_j^{(1)}\right)$$

$$\begin{aligned}
= \frac{1}{(n(n-1)h_H(\mathbb{E}[W_{12}]))^2}&\Big[n(n-1)\mathbb{E}[W_{12}^2(H_2^{(1)})^2] + n(n-1)\mathbb{E}[W_{12}W_{21}H_2^{(1)}H_1^{(1)}]\\
&+n(n-1)(n-2)\mathbb{E}[W_{12}W_{13}H_2^{(1)}H_3^{(1)}] + n(n-1)(n-2)\mathbb{E}[W_{12}W_{23}H_2^{(1)}H_3^{(1)}]\\
&+n(n-1)(n-2)\mathbb{E}[W_{12}W_{31}H_2^{(1)}H_1^{(1)}] + n(n-1)(n-2)\mathbb{E}[W_{12}W_{32}(H_2^{(1)})^2]\\
&-n(n-1)(4n-6)\mathbb{E}[W_{12}H_2^{(1)}]^2\Big].
\end{aligned} \tag{17}$$

We get, after some direct calculations

$$\begin{cases} \mathbb{E}[W_{12}^2 H_2^{(1)}] = O(h_K^4 h_H \phi_x^2(h_K)), \quad \mathbb{E}[W_{12}W_{21}H_2^{(1)}H_1^{(1)}] = O(h_K^4 h_H^2 \phi_x^2(h_K)), \\ \mathbb{E}[W_{12}W_{13}H_2^{(1)}H_3^{(1)}] = \mathbb{E}[W_{12}W_{31}H_2^{(1)}H_1^{(1)}] = \mathbb{E}[W_{12}W_{23}H_2^{(1)}H_3^{(1)}] = O(h_K^4 h_H^2 \phi_x^3(h_K)), \\ \mathbb{E}[W_{12}W_{32}(H_2^{(1)})^2] = \mathbb{E}^2[\beta_1^2 K_1]\mathbb{E}[K_1^2(H_1^{(1)})^2] + o(h_K^4 h_H \phi_x^3(h_K)). \end{cases}$$

Clearly, the latter term in the last cases is the leading one, and can be evaluated in (17) by using

$$\frac{(n-2)}{n(n-1)(h_H\mathbb{E}[W_{12}])^2}\mathbb{E}^2[\beta_1^2 K_1]\mathbb{E}[K_1^2(H_1^{(1)})^2]$$

So, after the same steps as in the previous Lemma, it suffices to write

$$Var\left(\widehat{f}_N^x(y)\right) = \frac{\mathbb{E}[K_1^2(H_1^{(1)})^2]}{n(h_H\mathbb{E}[K_1])^2} + o\left(\frac{1}{nh_H\phi_x(h_K)}\right). \tag{18}$$

Thus, by the change of variables $t = h_H^{-1}(y-z)$, we get

$$\mathbb{E}[K_1^2(H_1^{(1)})^2] = \mathbb{E}[K_1^2\mathbb{E}((H_1^{(1)})^2|X_1)]$$

and

$$\mathbb{E}((H_1^{(1)})^2|X_1) = h_H\int (H^{(1)})^2(t) f^{X_1}(y-h_H t)\mathrm{d}t.$$

Then, by Taylor's expansion of order 1 of $f^{X_1}(\cdot)$ we obtain

$$f^{X_1}(y-h_H t) = f^{X_1}(y) + O(h_H) = f^{X_1}(y) + o(1).$$

Now, it follows from (18) that:

$$\mathbb{E}[K_1^2(H_1^{(1)})^2] = h_H\int (H^{(1)})^2(t)\mathrm{d}t\,\mathbb{E}\left[K_1^2 f^X(y)\right] + o(h_H\mathbb{E}[K_1^2]).$$

Again, by the same steps in proof of Lemma 1, we get

$$\mathbb{E}\left[K_1^2 f^{X_1}(y)\right] = f^x(y)\mathbb{E}[K_1^2] + o(\mathbb{E}[K_1^2])$$

which implies:

$$\mathbb{E}[K_1^2(H_1^{(1)})^2] = h_H f^x(y)\mathbb{E}[K_1^2]\int (H^{(1)})^2(t)\mathrm{d}t + o(h_H\mathbb{E}[K_1^2]). \tag{19}$$

Consequently, we obtain from (16), (18) and (19), that

$$Var\left(\widehat{f}_N^x(y)\right) = \frac{f^x(y)}{nh_H\phi_x(h_K)}\left(\int H^{(1)}(t)^2\mathrm{d}t\right)\left[\frac{\left(K^2(1) - \int_{-1}^1 (K^2(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1) - \int_{-1}^1 (K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)^2}\right]$$

$$+o\left(\frac{1}{nh_H\phi_x(h_K)}\right).\blacksquare$$

Proof of Lemma 3. By simple computations, we have

$$\begin{aligned}
Cov\left(\widehat{f}_N^x(y), \widehat{f}_D(x)\right) &= \frac{1}{(n(n-1)h_H(\mathbb{E}[W_{12}]))^2} Cov\Big(\sum_{1\le i\neq j\le n} W_{ij}H_j^{(1)}, \sum_{1\le i'\neq j'\le n} W_{i'j'}\Big) \\
&= \frac{1}{(n(n-1)h_H(\mathbb{E}[W_{12}]))^2}\Big[n(n-1)\mathbb{E}[W_{12}^2H_1^{(1)}] + n(n-1)\mathbb{E}[W_{12}W_{21}H_2^{(1)}] \\
&\quad +n(n-1)(n-2)\mathbb{E}[W_{12}W_{13}H_2^{(1)}] + n(n-1)(n-2)\mathbb{E}[W_{12}W_{23}H_2^{(1)}] \\
&\quad +n(n-1)(n-2)\mathbb{E}[W_{12}W_{31}H_2^{(1)}] + n(n-1)(n-2)\mathbb{E}[W_{12}W_{32}H_2^{(1)}] \\
&\quad -n(n-1)(4n-6)(\mathbb{E}[W_{12}H_2^{(1)}]\mathbb{E}[W_{12}]\Big].
\end{aligned}$$

By direct manipulations, we get

$$\begin{cases}
\mathbb{E}[W_{12}^2H_2^{(1)}] = \mathbb{E}[W_{12}W_{21}H_2^{(1)}] = O(h_K^4 h_H \phi_x^2(h_K)), \\
\mathbb{E}[W_{12}W_{13}H_2^{(1)}] = \mathbb{E}[W_{12}W_{31}H_2^{(1)}] = O(h_K^4 h_H \phi_x^3(h_K)), \\
\mathbb{E}[W_{12}W_{23}H_2^{(1)}] = \mathbb{E}[W_{12}W_{32}H_2^{(1)}] = O(h_K^4 h_H \phi_x^3(h_K)).
\end{cases}$$

Since $\mathbb{E}[W_{12}] = O(h_K^2\phi_x^2(h_K))$, we obtain

$$Cov\left(\widehat{f}_N^x(y), \widehat{f}_D(x)\right) = O\left(\frac{1}{n\phi_x(h_K)}\right). \blacksquare$$

Proof of Lemma 4. The demonstration of this result follows the lines of the proof of the previous lemma, step by step, by replacing $H^{(1)}$ by 1. Thus,

$$\begin{aligned}
Var(\widehat{f}_D^x) &= \frac{1}{(n(n-1)\mathbb{E}[W_{12}])^2} Var\left(\sum_{1\le i\neq j\le n} W_{ij}\right) \\
&= \frac{1}{(n(n-1)\mathbb{E}[W_{12}]))^2}\Big(n(n-1)\mathbb{E}[W_{12}^2] + n(n-1)\mathbb{E}[W_{12}W_{21}] \\
&\quad +n(n-1)(n-2)\mathbb{E}[W_{12}W_{13}] + n(n-1)(n-2)\mathbb{E}[W_{12}W_{23}] \\
&\quad +n(n-1)(n-2)\mathbb{E}[W_{12}W_{31}] + n(n-1)(n-2)\mathbb{E}[W_{12}W_{32}] \\
&\quad -n(n-1)(4n-6)(\mathbb{E}[W_{12}])^2\Big).
\end{aligned}$$

Still by straightforward manipulations, we get

$$\begin{cases}
\mathbb{E}[W_{12}^2] = \mathbb{E}[W_{12}W_{21}] = O(h_K^4\phi_x^2(h_K)), \\
\mathbb{E}[W_{12}W_{13}] = \mathbb{E}[W_{12}W_{31}] = O(h_K^4\phi_x^3(h_K)), \\
\mathbb{E}[W_{12}W_{23}] = \mathbb{E}[W_{12}W_{32}] = O(h_K^4\phi_x^3(h_K)).
\end{cases}$$

So, we have

$$Var\left(\widehat{f}_D^x\right) = O\left(\frac{1}{n\phi_x(h_K)}\right). \blacksquare$$

Proof of Theorem 3. The proof of this theorem is based on the same techniques as in the proof of Theorem 2, where

$$\mathbb{E}\left[\widehat{F}^x(y) - F^x(y)\right]^2 = \left[\mathbb{E}\left[\widehat{F}^x(y)\right] - F^x(y)\right]^2 + Var\left[\widehat{F}^x(y)\right]$$

and to simplify the bias and the variance of the second term in the right equality, we use the results of Ferraty et al. (2007), to obtain

$$\mathbb{E}\big[\widehat{F}^x(y)\big] - F^x(y) = \Big(\mathbb{E}[\widehat{F}_N^x(y)] - F^x(y)\Big) + \frac{\mathbb{E}[\widehat{F}_N^x(y)(\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x])]}{(\mathbb{E}[\widehat{f}_D^x])^2} + \frac{\mathbb{E}[\widehat{F}^x(y)(\widehat{f}_D^x - \mathbb{E}[\widehat{f}_D^x])^2]}{(\mathbb{E}[\widehat{f}_D^x])^2}$$

and

$$\begin{aligned} Var\big[\widehat{F}^x(y)\big] &= Var\left(\widehat{F}_N^x(y)\right) - 4\left(\mathbb{E}[\widehat{F}_N^x(y)]\right) Cov\left(\widehat{F}_N^x(y), \widehat{f}_D(x)\right) \\ &+3\big(\mathbb{E}[\widehat{F}_N^x(y)]\big)^2 Var\left(\widehat{f}_D(x)\right) + o\left(\frac{1}{n\phi(h_K)}\right). \blacksquare \end{aligned}$$

Proof of Lemma 5. Concerning the quantities $\mathbb{E}[\widehat{F}_N^x(y)]$, we use an integration by part to arrive at

$$\mathbb{E}[\widehat{F}_N^x(y)] = \frac{1}{\mathbb{E}[W_{12}]}\mathbb{E}[W_{12}\mathbb{E}[H_2|X_2]] \text{ with } \mathbb{E}[H_2|X_2] = \int H_2^{(1)}(t)F^{X_2}(y - h_H t)\mathrm{d}t.$$

Then, the same steps used in studying $\mathbb{E}[\widehat{f}_N^x(y)]$ can be re-used to prove that

$$\begin{aligned} \mathbb{E}\left[\widehat{F}_N^x(y)\right] &= F^x(y) + \frac{h_H^2}{2}\frac{\partial^2 F^x(y)}{\partial y^2}\int t^2 H_2^{(1)}(t)\mathrm{d}t + o(h_H^2) \\ &+ \frac{h_K^2}{2}\Psi_{0,0}^{(2)}(0)\frac{\left(K(1) - \int_{-1}^1 (u^2 K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1) - \int_{-1}^1 K^{(1)}(u)\chi_x(u)\mathrm{d}u\right)} + o(h_K^2). \blacksquare \end{aligned}$$

Proof of Lemma 6. It clear that

$$\begin{aligned} Var[\widehat{F}_N^x(y)] &= \frac{1}{(n(n-1)h_H(\mathbb{E}[W_{12}]))^2}\Big[n(n-1)\mathbb{E}[W_{12}^2(H_2)^2] + n(n-1)\mathbb{E}[W_{12}W_{21}H_2H_1] \\ &+n(n-1)(n-2)\mathbb{E}[W_{12}W_{13}H_2H_3] + n(n-1)(n-2)\mathbb{E}[W_{12}W_{23}H_2H_3] \\ &+n(n-1)(n-2)\mathbb{E}[W_{12}W_{31}H_2H_1] + n(n-1)(n-2)\mathbb{E}[W_{12}W_{32}(H_2)^2] \\ &-n(n-1)(4n-6)\mathbb{E}[W_{12}H_2]^2\Big]. \end{aligned} \tag{20}$$

For these terms, we use the same steps used in Lemma 1 and we get

$$\begin{cases} \mathbb{E}[W_{12}^2H_2^2] = O(h_K^4\phi_x^2(h_K)), \mathbb{E}[W_{12}W_{21}H_1H_2] = O(h_K^4\phi_x^2(h_K)), \\ \mathbb{E}[W_{12}W_{13}H_2H_3] = (F^x(y))^2\mathbb{E}[\beta_1^4K_1^2]\mathbb{E}^2[K_1] + o(h_K^4\phi_x^3(h_K)), \\ \mathbb{E}[W_{12}W_{23}H_2H_3] = (F^x(y))^2\mathbb{E}[\beta_1^2K_1]\mathbb{E}[\beta_1^2K_1^2]\mathbb{E}[K_1] + o(h_K^4\phi_x^3(h_K)), \\ \mathbb{E}[W_{12}W_{31}H_2H_1] = (F^x(y))^2\mathbb{E}[\beta_1^2K_1]\mathbb{E}[\beta_1^2K_1^2]\mathbb{E}[K_1] + o(h_K^4\phi_x^3(h_K)), \\ \mathbb{E}[W_{12}W_{32}H_2^2] = F^x(y)\mathbb{E}^2[\beta_1^2K_1]\mathbb{E}[K_1^2] + o(h_K^4\phi_x^3(h_K)), \\ \mathbb{E}[W_{12}H_1] = O(h_K^2\phi_x^2(h_K)). \end{cases} \tag{21}$$

Hence, it follows from (20) and (21):

$$Var[\widehat{F}_N^x(y)] = \frac{F^x(y)(1-F^x(y))}{\mathbb{E}[K_1^2]}(\mathbb{E}[K_1])^2 + o\left(\frac{1}{n\phi_x(h_K)}\right).$$

Finally,

$$Var[\widehat{F}_N^x(y)] = \frac{F^x(y)(1-F^x(y))}{n\phi_x(h_K)}\left[\frac{\left(K^2(1) - \int_{-1}^1 (K^2(u))^{(1)}\chi_x(u)\mathrm{d}u\right)}{\left(K(1) - \int_{-1}^1 (K(u))^{(1)}\chi_x(u)\mathrm{d}u\right)^2}\right] + o\left(\frac{1}{n\phi_x(h_K)}\right). \blacksquare$$

Proof of Lemma 7. Both assertions of this lemma are direct consequences of Lemma 3.■

5. Conclusion and Perspectives

We presented in this paper the leading term of the mean square error of the estimator of the conditional hazard by the local linear approach. In terms of mean squared error our estimator performs competitively in comparison to existing estimators for the conditional hazard function. Our theoretical and practical studies confirm the superiority of the linear local approach over the classical kernel approach. From a theoretical point of view, there are interesting prospects. It would be very important in the next future to study the asymptotic normality of our estimator to make statistical tests. The kNN method is an alternative smoothing approach that offers an adaptive estimator. The very important feature of this method is that it allows the construction of a neighbourhood adapted to the local structure of the data. So, It would be also of interest to study the asymptotic properties of the kNN estimator of the conditional hazard function. This will be considered in future works.

Acknowledgements. The authors wish to thank the anonymous reviewers and the Editor in chief for insightful comments and suggestions that greatly helped to improve this paper.

References

Baìllo, A. and Grané, A. (2009). Local linear regression for functional predictor and scalar response. *Journal of Multivariate Analysis.* **100**, 102-111.

Barrientos-Marin, J., Ferraty, F. and Vieu, P. (2010). Locally modelled regression and functional data. *Journal of Nonparametric Statistics.* **22(5)**, 617-632.

Berlinet, A., Elamine, A. and Mas, A. (2011). Local linear regression for functional data. *Annals of the Institute of Statistical Mathematics.* **63(5)**, 1047-1075.

Boj, E., Delicado, P. and Fortiana, J. (2010). Distance-based local linear regression for functional predictors. *Computational Statistics and Data Analysis.* **54**, 429-437.

Bosq, D. and Lecoutre, J. P. (1987). *Théorie de l'estimation fonctionnelle.* Economica.

Bosq, D. (2000). *Linear Processes in Function Spaces: Theory and applications.* Lecture Notes in Statistics. New York, Springer.

Demongeot, J., Laksaci, A., Madani, F. and Rachdi, M. (2010). Local linear estimation of the conditional density for functional data. *C. R., Math., Acad. Sci. Paris.* **348**, 931-934.

Demongeot, J., Laksaci, A., Rachdi, M. and Rahmani, S. (2014). On the Local Linear Modelization of the Conditional Distribution for Functional Data. *Sankhya : The Indian Journal of Statistics.* **76**(2), 328-355.

El methni, M. and Rachdi, M. (2011). Local weighted average estimation of the regression operator for functional data. *Communications in Statistics-Theory and Methods.* **40**, 3141-3153.

Fan, J. (1992). Design-adaptative nonparametric regression. *Journal of the American Statistical association*. **87**, 998-1004.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and its Applications*. Monographs on Statistics and Applied Probability. **66**, Chapman& Hall.

Fan, J. and Yao, Q. (2003). *Non linear time series. Nonparametric and parametric methods.* Springer Series in Statistics. Springer-Verlag, New York.

Ferraty, F. and Vieu, P. (2002). Nonparametric models for functional data, with application in regression, time series prediction and curve discrimination, presented at the International Conference on Advances and Trends in Nonparametric Statistics. Crete, Greece.

Ferraty, F. and Vieu, P. (2006). Nonparametric Functional Data Analysis. *Springer Series in Statistics*. New York, Springer.

Ferraty, F., Mas, A. and Vieu, P. (2007). Non-paramétrique regression on functional data: infernce and practical aspects. *Aust. N. Z. J. Stat.* **49(3)**, 207-286.

Ferraty, F., Rabhi, A. and Vieu, P. (2008). Estimation non-paramétrique de la fonction de hasard avec variable explicative fonctionnelle. *Revue de Mathématiques Pures et Appliquées*. **53**, 1-18.

Massim, I. and Mechab, B. (2016). Local linear estimation of the conditional hazard function. *International Journal of Statistics and Economics*. **17**, 1-11.

Quintela-Del-Río, A. (2008). Hazard function given a functional variable: Nonparametric estimation under strong mixing conditions. *Journal of Nonparametric Statistics*. **20**, 413-430.

Rabhi, A., Benaissa, S., Hamel, E. and Mechab, B. (2013). Mean square error of the estimator of the conditional hazard function. *Appl. Math. (Warsaw)*. **(40)4**, 405-420.

Rachdi, M., Laksaci, A., Demongeot, J., Abdali, A. and Madani, F. (2014). Theoretical and practical aspects of the quadratic error in the local linear estimation of the conditional density for functional data. *Computational Statistics and Data Analysis*. **73**, 53-68.

Ramsay, J.O. and Silverman, B.W. (2005). *Functional data analysis.* Springer Series in Statistics, New York.

Roussas, G. (1989). Hazard rate estimation under dependence conditions. *Journal of Statistical Planning and Inference*. **22**, 81-93.

Spierdijk, L. (2008). Non-parametric conditional hazard rate estimation: A local linear approach. *Comput. Stat. Data Anal.* 52, 2419-2434.

Watson, G.S. and Leadbetter, M.R. (1964). Hazard analysis. *Sankhyia.* **26**, 101-116.